(4)

AD_____

USE OF RECOMBINANT DNA TECHNIQUES FOR THE PRODUCTION OF A MORE EFFECTIVE ANTHRAX VACCINE

ANNUAL REPORT

Donald L. Robertson

August 31, 1988

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, MD 21701-5012

Contract No. DAMD17-85-C-5167

Brigham Young University
Provo, UT 84602

Approved for Public Release;

Distribution unlimited

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for Public Release; Distribution Unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Brigham Young University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Provo, Utah 84602	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION U.S. Army Medical Research & Development Command 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAMD17-85-C-5167
8c. ADDRESS (City, State, and ZIP Code) Ft. Detrick, Frederick, MD 21701-5012	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 63763A; PROJECT NO. 3M2 637630807; TASK NO. AA; WORK UNIT ACCESSION NO. 041

11. TITLE (Include Security Classification)
(u) Use of Recombinant DNA Techniques for the Production of a More Effective Anthrax Vaccine

12. PERSONAL AUTHOR(S)
Donald L. Robertson, Ph.D.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual	FROM 6/1/87 TO 7/31/88	1988 Aug. 31	34

16. SUPPLEMENTARY NOTATION

COSATI CODES		
FIELD	GROUP	SUB-GROUP
06	13	
06	01	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
RA 1,

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We have continued to characterize the edema factor (EF) and lethal factor (LF) genes of Bacillus anthracis. Using site-specific mutagenesis, we have identified the ATP binding site of EF, which is a calmodulin-dependent adenylate cyclase. We have also identified a potential calmodulin binding site. In addition, we have shown that carboxyl terminus of EF shares at least three highly conserved amino acid domains with the calmodulin-dependent adenylate cyclase of Bordetella pertussis.

We have also determined the DNA sequence of the LF gene. This gene contains many features in common with the EF and PA genes, including a strong ribosome binding site and a long signal peptide (33 amino acids) which is apparently removed during secretion. In addition, the amino terminal 300 amino acids of LF shares extensive homology with the amino terminus of EF. This shared homology

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Mary Frances Bostian	22b. TELEPHONE (Include Area Code) (301) 663-7325
	22c. OFFICE SYMBOL SGRD-RMI-S

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

is probably part of their binding domains for associating with PA since EF and LF each bind to PA.

Each of the anthrax toxin genes (cya [EF], pag [PA] and lef [LF] have been expressed in Bacillus subtilis. In addition, expression plasmids have been constructed for regulated high level expression of these genes. We have also fused the EF coding region to the PA promoter and then expressed EF in B. subtilis using this plasmid construction. This plasmid has also been transferred into B. anthracis for high level, regulated expression. Expression of the individual toxin genes in B. subtilis should provide a safe bacterial host for production of large quantities of the B. anthracis toxin protiens.

SUMMARY OF RESEARCH

The overall goal of the present research is to construct a safe, effective human anthrax vaccine using recombinant DNA techniques. These studies are broken down into three phases:

Phase I. Isolation and characterization of the *Bacillus anthracis* toxin genes for protective antigen (PA), lethal factor (LF) and edema factor (EF). The individual toxin genes will be cloned in expression vectors for large scale production of toxin proteins using *E. coli* and *B. subtilis*. These experiments should provide enhanced production of the different toxin components which are made in low levels in *E. coli*.

Phase II. Generation of mutant toxin proteins from cloned toxin genes defined in Phase I. Mutations derived from deletion analysis or site-specific mutagenesis of the cloned toxin genes will be generated using in vitro manipulations of the recombinant plasmid DNAs. Mutations of potential use for vaccine construction will be identified as those which are non-toxic but still immunologically active and protective.

Phase III. Insertion of mutant genes back into *B. anthracis* with the selective removal of wild-type genes. Then, testing of these mutant strains will be performed in animals, such as the mouse or guinea pig. The research outlined in this annual report describes the cloning and characterization of the individual *B. anthracis* toxin genes. These genes are being expressed in *B. subtilis* and *E. coli* and are being specifically mutated to generate mutant derivatives which lack biochemical activity but maintain immunological properties. In addition, a physical characterization of the *B. anthracis* plasmids with regard to size, genetic complexity, GC% and restriction enzyme mapping is also described.

FOREWORD

The investigators (Principal Investigator and Graduate Students) have abided by the National Institutes of Health Guidelines for Research Involving Recombinant DNA Molecules (May, 1986) Supplemental guidelines pertaining to the subcloning of the individual *B. anthracis* toxin genes in sporulation competent *B. subtilis* was approved by the NIH committee on toxins March 13, 1986. All recombinant DNA research has also been registered with and approved by the Brigham Young University Institutional Biosafety Committee.

TABLE OF CONTENTS

BACKGROUND

As discussed in the summary, the goal of the experiments performed in this laboratory is to develop a more effective human anthrax vaccine for the protection of U.S. Army troops using recombinant DNA techniques. The current human anthrax vaccine consists of alum-precipitated supernatant material from fermenter cultures of *B. anthracis* which consists predominantly of PA (protective antigen) (1). Unfortunately, this vaccine may not be effective against all strains of *B. anthracis* since several virulent strains have been classified as "vaccine resistant" with regard to this human vaccine (2). Clearly, an effective vaccine must afford immunological protection against all strains of *B. anthracis* and against all forms of infection, including aerosol.

Virulent strains of *B. anthracis* contain two different plasmids. The toxin plasmid (pXO1) is necessary for expression of the three toxin proteins (3,4) and the capsule plasmid (pXO2) (5,6) is necessary for production of the poly-D-glutamic acid capsule (5,7). In order to be able to insert mutant toxin genes back into *B. anthracis* for the production of a safe vaccine strain it has been necessary to characterize these plasmids. Studies designed to physically characterize these plasmids have included buoyant density centrifugation, DNA melting analysis and restriction endonuclease mapping of these DNAs. These characterizations should be helpful in generating recombinant vaccine strains of *B. anthracis* and in understanding the physical organization of these DNAs.

Each of the anthrax toxin genes are cloned (4,8,9). The PA and EF genes have been sequenced (13,15). Experiments which are aimed at expressing these toxin genes in large quantities in *E. coli*, *B. subtilis* and *B. anthracis* are in progress. In addition, we are mutating the different toxin genes

in order to generate mutant toxin proteins which are still immunogenic but biochemically non-functional to be used in vaccine development.

RESULTS

Restriction maps of pXO1 and pXO2. The restriction maps for pXO1 and pXO2 (see Figures 1 and 2) are essentially completed for enzymes which cleave a few times, such as *Pst*I, *Bam*HI, *Cla*I, *Sst*I, *Bgl*II and *Pvu*II. Experiments to map the more frequent cutting enzymes, such as *Eco*RI and *Hind*III, are presently being completed. We have generated recombinant DNA libraries for pXO1 and pXO2 in bacteriophage λ as well as in plasmids in order to generate a complete map for the most commonly used restriction enzymes. A detailed restriction enzyme map of the LF and PA gene regions on pXO1 is shown in Figure 3.

In a final effort to map pXO1 and pXO2, we are identifying the number and location of the different RNA transcripts from these plasmids. This research project involves the identification of the different promoters and the RNAs made from them. Basically, we are cleaving pXO1 and pXO2 with an enzyme which cleaves these DNAs many times, such as *Mbo*I or *Sau*3A, generating DNA fragments which can ligate to *Bam*HI cleaved plasmids. Using *B. subtilis* plasmids which have been cleaved with *Bam*HI located prior to a promoterless chloramphenicol resistance gene (10), we will insert the pXO1 or pXO2 DNA fragments into these promoter identification plasmids. After transformation of these recombinant plasmids into *B. subtilis*, we will identify bacteria which are now resistant to chloramphenicol. These plasmids will contain a functional promoter (derived from pXO1 or pXO2) driving the transcription of the chloramphenicol resistance gene. The recombinant DNA inserts prepared from these promoter expression plasmids

will then be mapped on pXO1 or pXO2. The size and direction of RNA transcription will also be determined. This procedure is very powerful and should allow us to identify and position most, if not all, of the functional promoters from the *B. anthracis* plasmids, assuming that all these promoters will also function in *B. subtilis*. However, with the recent discovery that we can transform *B. anthracis* using electroporation, we will also be able to transfer these promoter plasmids to *B. anthracis* for promoter identification directly in the parent organism.

Characterization of the edema factor gene (*cya*). The edema factor is a calmodulin-dependent adenylate cyclase (11,12). We have successfully cloned and sequenced the EF gene (*cya*) and the DNA sequence (13) was reported in the previous annual report. A paper describing the cloning and expression of EF in *E. coli* has been published (9) and a manuscript describing the DNA sequence and its deduced amino acid sequence has been submitted and should soon be accepted by Gene (13). We have included the complete EF amino acid sequence, deduced from its DNA sequence, in Appendix I.

Several interesting structural features for EF are part of its deduced amino acid sequence. (1) EF apparently contains a 33 amino acid signal peptide which conforms to known *Bacillus* leader sequences in that it starts with charged (mostly positive) and hydrophilic residues (amino acids 1-10), followed by a central core of hydrophobic amino acids (residues 11-23) and then several hydrophilic residues (amino acids 24-33) prior to the start of the mature protein. Proteolytic cleavage apparently occurs at an Ala-Met peptide bond, near the start of a proposed α-helix (see Figure 4A), consistent with signal processing after an Ala or Gly in bacilli (14). A 29 amino acid leader sequence was also found for PA (15) which would likely contain a similar secondary structure (shown in Figure 3A). Likewise, a signal

peptide of 33 amino acids would be present in the LF-precursor molecule (Figure 3A). Figure 4B shows a comparison between the amino acid sequences near the ends of the EF, PA and LF signal peptides and the apparent position of proteolytic cleavage. Similar amino acids at the end of the signal peptide may be required for signal peptidase recognition or for secretion. (ii) A very strong *Bacillus* ribosome binding site immediately upstream from the start of the EF protein coding region is present (AAAGGAGGT) which is similar to the PA and LF ribosome binding sites (both of these genes have a ribosome binding site sequence of AAAGGAG). (iii) Amino acid residues 347 to 355 of the EF-precursor protein contains the sequence Gly-x-x-x-x-Gly-Lys-Ser (where x=any amino acid) which is a perfect match to a consensus sequence present in prokaryotic and eukaryotic ATP and GTP binding proteins (16). The Lys residue is part of the ATP binding sites of these proteins and appears to be part of the EF ATP binding site as well. That is, using site-specific mutagenesis procedures, we have replaced this Lys within EF with an Asn and cyclase activity was reduced 90-95% (unpublished data of author). (iv) We have also identified a domain in EF which could represent its putative calmodulin-binding site. As described in the EF sequencing paper (13), calmodulin-binding proteins often contain an α-helical region with charged or hydrophilic residues on one side and hydrophobic residues on the other. Such an amphiphilic helical region is present in EF located between amino acid residues 313-323 of the EF-precursor (see Appendix IV). Interestingly, this putative calmodulin-binding site is conserved in the *B. pertussis* adenylate cyclase as well (17,18). (v) No homology between the EF gene or its deduced EF amino acid sequence was observed with either the *E. coli* or yeast adenylate cyclases. However, there is at least three regions of homology in the amino acid sequence between EF and the *B.*

pertussis calmodulin-dependent adenylate cyclase. A section describing this homology is including below.

Characterization of the LF gene (*lef*). We have determined the entire DNA sequence for the *B. anthracis* LF gene (*lef*). We easily identified the start of the LF gene since the first 15 amino acids of the mature LF was previously determined by Dr. J. Schmidt (USAMRIID). The LF DNA sequence and the deduced amino acid sequence are shown in Appendix II. The LF gene contains a good ribosome binding site (AAAGGAG) which is identical to the proposed PA gene ribosome binding site. The LF-precursor apparently contains a 33 amino acid signal sequence (see Figure 4A) which is removed during secretion. This signal sequence conforms to consensus *Bacillus* leader peptides (and to the EF and PA signal peptides) in that it starts with a polar or charged region followed by 23 non-polar, hydrophobic amino acid residues. After this 33 amino acid leader peptide, the next 16 amino acids correspond exactly to the LF amino acid sequence determined by Dr. Jim Schmidt (USAMRIID), except for one amino acid. Amino acid position +10 of the mature protein (+43 of LF-precursor) is a His (based on the DNA sequence) whereas it was previously reported to be a Lys (based on LF protein sequencing). Interestingly, there is a single Cys in the LF leader, although no Cys residues are in the mature protein. The entire protein sequence of LF is also shown in Appendix III. Lastly, there appears to be extensive amino acid homology between LF and EF in the first 300 amino acids of these proteins. We have detected 10 closely elated domains and three of these highly conserved domains are underlined (and labelled #1, #2 and #3) in Appendix I and Appendix III. These homologous regions could represent PA binding domains. Since most of these domains are highly charged, interactions with PA may occur through a series of electrostatic interactions.

Transcription start sites for the anthrax toxin genes. Using radiolabeled oligonucleotides specific for each of the different toxin genes, we have attempted to determine the start site for transcription. Using mRNA (isolated from *B. anthracis* Sterne) as template, each oligonucleotide was used to prime DNA synthesis (using reverse transcriptase) towards the 5'-end of the respective toxin mRNA. This newly synthesized radioactive DNA was denatured and electrophoresed on a denaturing polyacrylamide gel. Using this approach, we have successfully identified the start sites for PA and LF gene transcription. The PA promoter is apparently located immediately upstream from the start of its coding region with transcription starting about 25 bases before the first start codon for PA translation (15). Likewise, the apparent start for LF gene transcription occurs 25 bases prior to the ATG start codon for LF translation (about nucleotide 456 in Appendix II). We have not yet been able to localize EF gene transcription. This failure is probably due to the low level of EF mRNA produced in *B. anthracis* which is at least 10-fold lower than either the PA or LF mRNA concentrations (unpublished data of author).

Site-specific mutagenesis of the EF gene. Using site-specific mutagenesis procedures, we have altered the EF gene in order to modify its enzyme activity and to construct EF expression vectors. First, the previously identified ATP binding domain in EF, which conforms to the consensus ATP binding site of other prokaryotic and eukaryotic ATP and GTP binding proteins (16), has a Lys residue which has been shown to be involved in ATP binding, was changed to an Asn in EF. The EF adenylate cyclase activity of this mutant, isolated from *E. coli*, was reduced about 90-95% indicating that this Lys is probably involved in ATP binding. However, since total activity was not abolished, other residues are probably also involved. Of particular interest,

is the presence of a His two residues prior to this Lys. This His is also conserved in the *B. pertussis* adenylate cyclase as discussed below (see also Appendix IV).

We have also removed the *BglII* cleavage site within the EF gene and inserted a new *BglII* recognition site immediately prior to the start of the protein coding sequence. In another experiment, we inserted a *BglII* cleavage site immediately downstream from the PA promoter so that we could fuse the PA promoter to the EF gene. This hybrid toxin gene, when inserted into pBS42 (19) and transformed into *B. subtilis*, expressed EF at a level at least as great as *B. anthracis* Sterne. We are in the process of determining the precise amount produced using an ELISA or Western blot. EF was secreted from *B. subtilis* and was enzymatically active in an adenylate cyclase assay. Since PA expression is regulated by bicarbonate (20) in *B. anthracis* (Dr. J. Bartkus, USAMRIID, personal communication), we are attempting to transfer this PA promoter-EF gene plasmid into *B. anthracis* by electroporation. Hopefully, this plasmid, when introduced into *B. anthracis* will produce regulated high levels of EF for purification and analysis. EF gene mutants can also be generated and transferred to *B. anthracis* using this plasmid construction.

Expression of toxin genes in *B. subtilis* and *B. anthracis*. In an effort to express the toxin genes in *B. subtilis* for secretion, we cloned each of the genes into *B. subtilis* plasmids. Initially, we expressed these genes by cloning them to a regulated promoter (in plasmid pSI-1) which also contains a strong ribosome binding site (21). For these constructions, we introduced unique *XbaI* cleavage sites prior to the start codons for the PA, EF and LF genes. Following cleavage with *XbaI* (which does not cleave within either the EF or PA genes), the entire toxin gene was ligated into

plasmid pSI-1. When transformed into *B. subtilis*, transcription of the inserted toxin genes was regulated by the *lac* repressor and IPTG (19,21). For these hybrid genes, the amount of PA produced was close to the amount produced by PA1 (22; unpublished data of author).

We also created a toxin expression plasmid using the T7 promoter cloned upstream from the toxin gene. In order to get expression in *B. subtilis*, we introduced into *B. subtilis* a cloned copy of the T7 RNA polymerase gene (23). These bacteria contain the T7 polymerase as part of an integration plasmid for regulated expression since the T7 gene was inserted into the regulatable promoter site of pSI-1. In order to select for cells containing this polymerase, we also included the erythomycin resistence gene from pE194, prior to integration into *B. subtilis* genomic DNA (24,25). *B. subtilis* containing this integration plamsid should express T7 RNA polymerase after the addition of IPTG. These cells will then be transformed with a replication competent plasmid containing one of the *B. anthracis* toxin genes (e.g., *cya*, *pag*, or *lef*) cloned downstream from the T7 promoter for gene expression. Although we have not yet tested these recombinant *B. subtilis*, these plasmid constructions express toxin in *E. coli* using the T7 polymerase. *B. subtilis* containing these plasmids should produce high level, regulated expression of the toxin genes in a safe bacterial host. Toxin protein is secreted and can be used for purification of individual toxin components.

Relationships between EF and the pertussis adenylate cyclase. *Bordetella pertussis*, the causative agent of whopping cough, secretes, among other virulence factors, a calmodulin-dependent adenylate cyclase. The adenylate cyclase appears to function independently of the pertussis toxin, but is a required virulence factor since strains which lack cyclase activity are avirulent (26). Glaser et al. (18) recently showed that the cyclase catalytic

domain is about 450 amino acids in length and is part of a larger precursor polypeptide of 1706 amino acids. With the anticipation that EF and the pertussis cyclase might be related, we performed a homology search between the entire EF (800 amino acids) and pertussis cyclase translational products (1706 amino acids). Three major regions of homology (labeled #1, #2 and #3 in Appendix IV) were observed which included the catalytic domain of the pertussis cyclase and the carboxyl terminal 500 amino acids of EF. This homology comparison is shown in Appendix IV. Domain #1 contains the consensus ATP binding site which is surrounded by highly conserved amino acids. This high degree of amino acid conservation indicates a close evolutionary relatedness for these two proteins. The putative calmodulin-binding site is conserved for these proteins and is shown in Appendix IV.

Restriction endonuclease cleavage maps for the anthrax toxin genes. Using the DNA sequences for the EF, PA and LF toxin genes, we have generated a set of restriction endonuclease cleavage maps for these genes. These are shown in Appendices V, VI and VII. These maps should be helpful to those researchers using DNA containing these genes.

CONCLUSIONS

It appears from the data reported here that Phase I, II and III of the original research proposal are essentially completed. Each of the anthrax toxin genes has been cloned and expressed in *E. coli* and to some extent in *B. subtilis* and *B. anthracis*. Since we have cloned each of the toxin genes and know their DNA sequences, we will be able to continue to study gene expression and to characterize the toxin proteins better. We will be able to generate toxin gene mutants for the construction of a safe vaccine and to elucidate the biochemical activities of these proteins. With the exception of putting the mutant genes back into *B. anthracis*, our research will allow us to construct a safe recombinant DNA derived anthrax vaccine.

LITERATURE CITED

1. Ristroph, J. H., and B. E. Ivins. 1983. Production of high levels of *Bacillus anthracis* toxin antigens in a new, defined culture medium. Infect. Immun. 39:483-486.

2. Little, S.F. and G.B. Knudson. 1986. Comparative efficacy of *Bacillus anthracis* live spore vaccine and protective antigen vaccine against anthrax in the guinea pig. Infect. Immun. 52:509-512.

3. Mikesell, P., B. E. Ivins, J. D. Ristroph, and T. M. Dreier. 1983. Evidence for plasmid-mediated toxin production in Bacillus anthracis. Infect. Immun. 39:371-376.

4. Vodkin, M. H., and S. H. Leppla. 1983. Cloning of the protective antigen gene of Bacillus anthracis. Cell 34:693-697.

5. Green, B.D, L. Battisti, T.M. Koehler, C.B. Thorne, and B.E. Ivins. 1985. Demonstration of a capsule plasmid in *B. anthracis*. Inf. Immun. 49:291-297.

6. Uchida, I., Sekizaki, T., Hashimoto, K., and Terakado, N. 1985. Association of the encapsulation of *Bacillus anthracis* with a 60 megadalton plasmid. J. of Gen. Micro. 131:363-367.

7. Zwartouw, H. T., and H. Smith. 1956. Polyglutamic acid from *Bacillus anthracis* grown in vivo: structure and aggressin activity. Biochem. J. 63:437-442.

8. Robertson, D. L., and S. H. Leppla. 1986. Molecular cloning and expression on *Escherichia coli* of the lethal factor gene of *Bacillus anthracis*. Gene 44:71-78.

9. Tippetts, M.T. and D.L. Robertson. 1988. Molecular cloning and expression of the Bacillus anthracis edema factor toxin gene: a calmodulin-dependent adenylate cyclase. J. Bacteriol. 170:2263-2266

10. Lovett, P.S. and S. Mongkolsuk. 1988. Promoter probe plasmids for Gram-positive Bacteria. *in* Vectors: A survey of molecular cloning vectors and their uses. ed. R.L. Rodriquez and D.T. Denhardt. Butterworths.

11. Leppla, S. H. 1982. Anthrax toxin edema factor: A bacterial adenylate cyclase that increases cyclic AMP concentrations in eukaryotic cells. Proc. Natl. Acad. Sci. USA 79:3162-3166.

12. Leppla, S. H. 1984. *Bacillus anthracis* calmodulin-dependent adenylate cyclase: Chemical and enzymatic properties and interactions with eucaryotic cells. *in* Advances in Cyclic Nucleotide and Protein Phosphorylation Research. ed. P. Greengard et al. Vol. 17, pp. 189-198.

13. Robertson, D.L., M.T. Tippetts and S.H. Leppla. 1988. Nucleotide sequence of the *Bacillus anthracis* edema factor gene (*cya*): A calmodulin-dependent adenylate cyclase. Submitted to Gene.

14. Pugsley, A.P. and Schwartz, M. 1985. Export and secretion of proteins by bacteria. **FEMS Microbiol. Rev.** 32:3-38.

15. Welkos, S.L., Lowe, J.R., Eden-McCuthan, F., Vodkin, M., Leppla, S.H. and J.J. Schmidt, J.J. 1988. Nucleotide sequence of the protective antigen gene of *Bacillus anthracis*. in press, **Gene.**

16. Higgins, C.F., Hiles, I.D., Salmond, G.P.C., Gill, D.R., Downie, J.A., Evans, I.J., Holland, I.B., Gray, L., Buckel, S.D., Bell, A.W. and Hermodson, M.A. 1986. A family of ATP-binding subunits coupled to many distinct biological processes in bacteria. **Nature (London).** 323:448-450.

17. Robertson, D.L. 1988. Relationships between the calmodulin-dependent adenylate cyclases produced by *Bacillus anthracis* and *Bordetella pertussis*. Submitted to **Infect. Immun.**

18. Glaser, P., Ladant, D., Sezer, O., Pichot, F., Ullmann, A. & Danchin, A. 1988. The calmodulin-sensitive adenylate cyclases of *Bordetella pertussis*: cloning and expression in *Escherichia coli*. **Mol. Microbiol.** 2:19-30.

19. Band, L. and D.J. Henner. 1984. *Bacillus subtilis* requires a "stringent" Shine-Dalgarno region for gene expression. **DNA** 3:17-21.

20. Sterne, M. 1939. The effects of different carbon dioxide concentrations on the growth of virulent anthrax strains. Pathogenicity and immunity tests on guinea pigs and sheep with anthrax variants derived from virulent strains. **Onderstepoort J. Vet. Sci. Anim. Husb.** 13:307-312.

21. Yansura, D.G. and D.J. Henner. 1984. Development of an inducible promoter for controlled gene expression in Bacillus subtilis. **Genetics and Biotechnology of Bacilli**, pp. 249-263.

22. Ivins, B.E. and S.L. Welkos. 1986. Cloning and expression of the *Bacillus anthracis* protective antigen gene in *Bacillus subtilis*. **Infect. Immun.** 54:537-542.

23. Studier, F.W. and B.A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. **J. Mol. Biol.** 189:113-130.

24. Saunders, C.W., B.J. Schmidt, M.S. Mirot, L.D. Thompson, and M.S. Guyer. 1984. Use of chromosomal integration in the establishment and expression of *blaZ*, a *Staphylococcus aureus* β-lactamase gene, in *Bacillus subtilis*. **J. Bacteriol** 157:718-726.

25. Saunders, C.W., C. D.B. Banner, S. R. Fahnestock, M. Lindberg, M.S. Mirot, C.S. Rhodes, C.F. Rudolph, B.J. Schmidt, L.D. Thompson, M. Uhlen, and M.S. Guyer. 1984. Use of chromosomal integration for the establishment of heterologous genes in *Bacillus subtilis*. *in* **Protein Transport and Secretion**, pp. 329-339.

26. Weiss, A.A., Hewlett, E.L., Myers, G.A. & Falkow, S. 1984. Pertussis toxin and extracytoplasmic adenylate cyclase as virulence factors in *Bordetella pertussis*. J. Infect. Dis. 150:219-222.

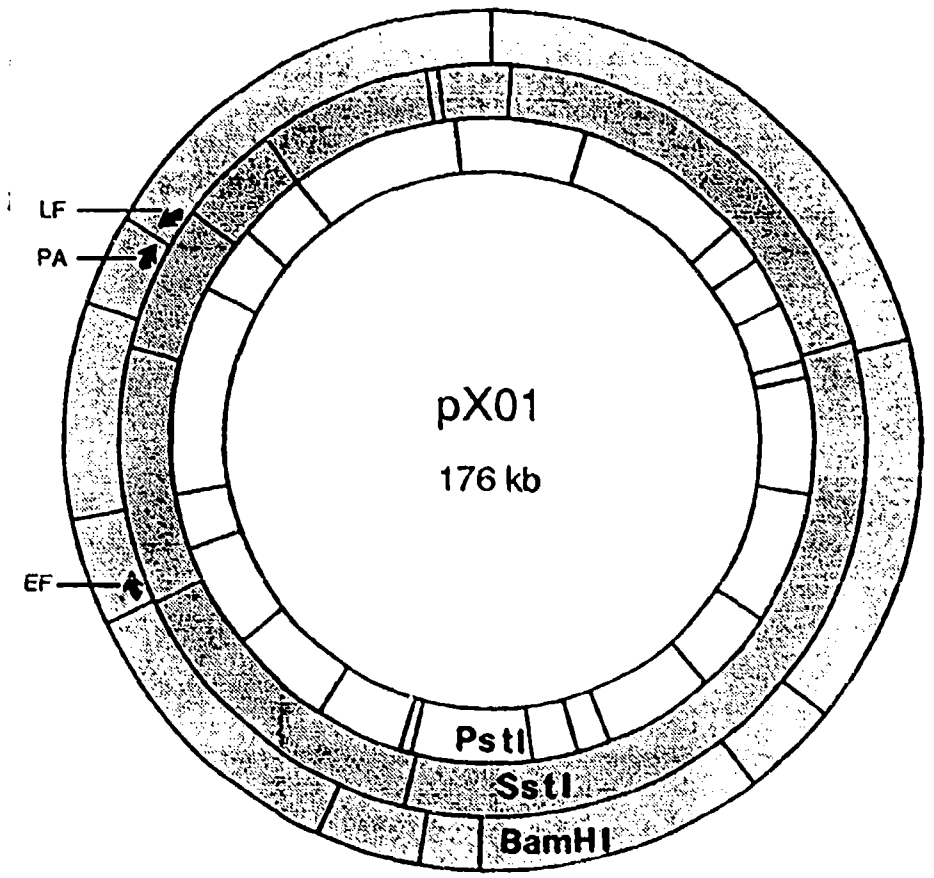


FIGURE 1. Restriction map of pXO1. The positions of the LF, PA and EF genes are depicted. The sizes of DNA fragments for each enzyme are not included due to the lack of space.

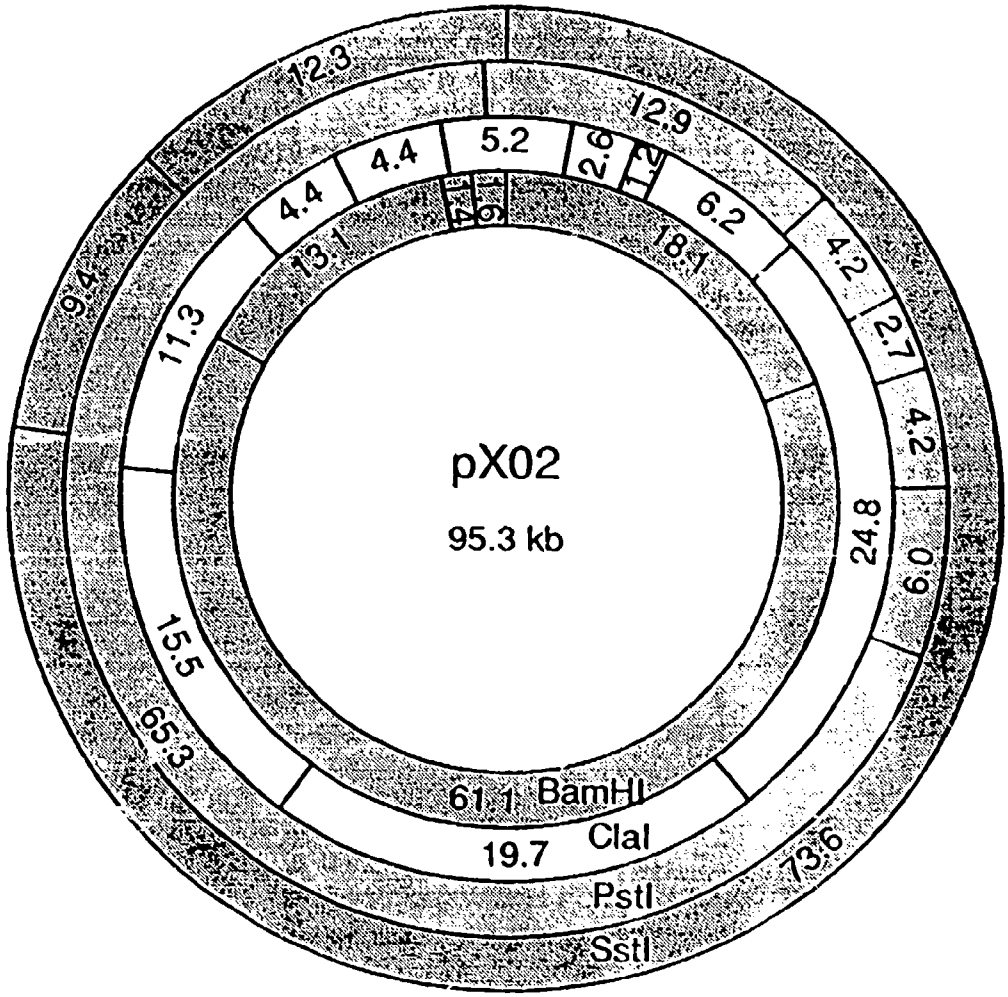
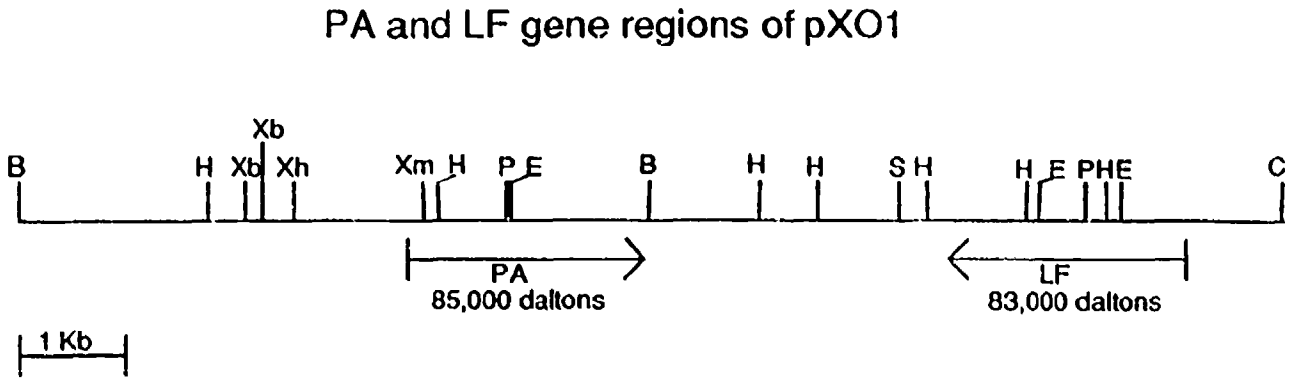


FIGURE 2. Restriction map of pX02.



B - *Bam* HI
H - *Hin* dIII
Xb - *Xba* I
Xh - *Xho* I
Xm - *Xmn* i
P - *Pst* I
E - *Eco* RI
S - *Sst* I
C - *Cla* I

FIGURE 3. Restriction map of the PA and LF gene regions on pXO1.

(A) The signal peptides (in bold) for EF, PA and LF are shown. The proposed secondary structure most likely to be assumed for the first 60 amino acids of each protein is shown (α=α-helix; β=β-sheet; t=β-turn; blank=random coil). The amino terminal amino acid, as determined by Dr. J. Schmidt (USAMRIID), for each mature toxin protein is also shown.

EF signal peptide

↓-start of mature EF

1 **MTRNKFIPNKFSIISFSVLL FAISSSQAIEVNAMNEHYTE SDIKRNHKTEKNKTEKEKFK** 60
αατtt ttββββββββββ ααα ααααααααααααα αααααααααααααααααααα

PA signal peptide

↓-start of mature PA

1 **MKKRKVLIPLMALSTILVSS TGNLEVIQAEVKQENRLLNE SESSSQGLLGYYFSDLNFQA** 60
αααααββββαααββββββ ααααααααααααα ttttttββββββββt ααα

LF signal peptide

↓-start of mature LF

1 **MNIKKEFIKVISMSCLVTAI TLSGPVFIPLVQGAGGHGDV GMHVKEKEKNKDENKRKDEE** 60
ααααααααααααtββββββ βt tββββββ α αααααααααααααααααααα

(B) The amino acid sequence at the end of the anthrax toxin signal peptides is shown. Cleavage occurs after Ala or Gly, consistent with known cleavages after bacilli signal peptides (14). Similar amino acids at the end of the signal peptides (denoted with a vertical bar [|]) probably represents signal peptidase recognition sequences. The numbers (-1 or +1) indicate the last amino acid of the signal peptide and the first amino acid of the mature toxin protein, respectively.

	-1 +1
EF signal peptide	Glu-Val-Asn-Ala--Met
	| | |
PA signal peptide	Val-Ile-Gln-Ala--Glu
	| | | |
LF signal peptide	Leu-Val-Gln-Gly--Ala

FIGURE 4. Anthrax toxin signal peptides.

APPENDIX I. EF amino acid sequence

(33 aa signal peptide) ↓-Start of mature EF (767 aa)

1 MTRNKFIPNKFSIISFSVLLFAISSSQAIEVNAMNEHYTESDIKRNHKTEKNKTEKEKFKDSINNLVKTE

71 FTNETLDKIQQTQDLLKKIPKDVLEIYSELGGEIYFTDIDLVEHKELQDLSEEEKNSMNSRGEKVPFASR

141 FVFEKKRETPKLIINIKDYAINSEQSKEVYYEIGKGISLDIISKDKSLDPEFLNLIKSLSDDSDSSDLLF
#1 #2

211 SQKFKEKLELNNKSIDINFIKENLTEFQHAFSLAFSYYFAPDHRTVLELYAPDMFEYMNKLEKGGFEKIS
#3

281 ESLKKEGVEKDRIDVLKGEKALKASGLVPEHADAFKKIARELNTYILFRPVNKLATNLIKSGVATKGLNE
(Potential calmodulin binding site)

351 HGKSSDWGPVAGYIPFDQDLSKKHGQQLAVEKGNLENKKSITEHEGEIGKIPLKLDHLRIEELKENGIIL
(Putative ATP binding site)

421 KGKKEIDNGKKYYLLESNNQVYEFRISDENNEVQYKTKEGKITVLGEKFNWRNIEVMAKNVEGVLKPLTA

491 DYDLFALAPSLTEIKKQIPTKRMDKVVNTPNSLEKQKGVTNLLIKYGIERKPDSTKGTLSNWQKQMLDRL

561 NEAVKYTGYTGGDVVNHCTEQDNEEFPEKDNEIFIINPEGEFILTKNWEMTGRFIEKNITGKDYLYYFNR

631 SYNKIAPGNKAYIEWTDPITKAKINTIPTSAEFIKNLSSIRRSSNVGVYKDSGDKDEFAKKESVKKIAGY

701 LSDYYNSANHIFSQEKKRKISIFRGIQAYNEIENVLKSKQIAPEYKNYFQYLKERITNQVQLLLTHQKSN

771 IEFKLLYKQILNFTENETDNFEVFQKIIDEK

****

The sequence contains 800 amino acids (M_r 92,464):

Ala (A)	32	Leu (L)	69
Arg (R)	22	Lys (K)	103
Asn (N)	61	Met (M)	9
Asp (D)	44	Phe (F)	40
Cys (C)	0	Pro (P)	23
Gln (Q)	27	Ser (S)	55
Glu (E)	82	Thr (T)	39
Gly (G)	40	Trp (W)	5
His (H)	13	Tyr (Y)	34
Ile (I)	68	Val (V)	34

Acidic	(Asp + Glu)	126
Basic	(Arg + Lys)	125
Aromatic	(Phe + Trp + Tyr)	79
Hydrophobic	(Aromatic + Ile + Leu + Met + Val)	259

APPENDIX II: Nucleotide Sequence of the LF gene.

```
      10        20        30        40        50        60        70        80        90
AAATTAGGATTTCGGTTATGTTTAGTATTTTTTTAAAATAATAGTATTAAATAGTGGAATGCAAATGATAAATGGGCTTTAAACAAAACT

     100       110       120       130       140       150       160       170       180
AATGAAATAATCTACAAATGGAATTTCTCCAGTTTTAGATTAAACCATACCAAAAAAATCACACTGTCAAGAAAAATGATAGAATCCCTA

     190       200       210       220       230       240       250       260       270
CACTAATTAACATAACCAAATTGGTAGTTATAGGTAGAAACTTATTTATTTCTATAATACCATGCAAAAAAGTAAATATTCTGTTCCATA

     280       290       300       310       320       330       340       350       360
GTATTTTAGTAAATTATTTAGCAAGTAAATTTTGGTGTATAAACAAAGTTTATCTTAATATAAAAAATTACTTTACTTTTATACAGATTA

     370       380       390       400       410       420       430       440       450
AAATGAAAAATTTTTTATGACAAGAAATATTGCCTTTAATTTATGAGGAAATAAGTAAAATTTTGTACATACTTTATTTTATTGTTGAAA

     460       470       480       490       500       510       520       530       540
TGTTCACTTATAAAAAAGGAGAGATTAAATATGAATATAAAAAAGAATTTATAAAAGTAATTAGTATGTCATGTTTAGTAACAGCAATT
          (r.b.s.)          MetAsnIleLysLysGluPheIleLysValIleSerMetSerCysLeuValThrAlaIle
                                  (33 amino acid signal peptide)

     550       560       570       580       590       600       610       620       630
ACTTTGAGTGGTCCCGTCTTTATCCCCCTTGTACAGGGGCGGCGCGGTCATGGTGATGTAGGTATGCACGTAAAAGAGAAAGAGAAAAAT
ThrLeuSerGlyProValPheIleProLeuValGlnGlyAlaGlyGlyHisGlyAspValGlyMetHisValLysGluLysGluLysAsn
                                  +1 of mature LF

     640       650       660       670       680       690       700       710       720
AAAGATGAGAATAACAGAAAAGATGAAGAACGAAATAAAACACAGGAAGAGCATTTAAAGGAAATCATGAAACACATTGTAAAAATAGAA
LysAspGluAsnLysArgLysAspGluGluArgAsnLysThrGlnGluGluHisLeuLysGluIleMetLysHisIleValLysIleGlu

     730       740       750       760       770       780       790       800       810
GTAAAAGGGGAGGAAGCTGTTAAAAAAGAGGCAGCAGAAAAGCTACTTGAGAAAGTACCATCTGATGTTTTAGAGATGTATAAAGCAATT
ValLysGlyGluGluAlaValLysLysGluAlaAlaGluLysLeuLeuGluLysValProSerAspValLeuGluMetTyrLysAlaIle

     820       830       840       850       860       870       880       890       900
GGAGGAAAGATATATATTGTGGATGGTGATATTACAAACATATATCTTTAGAAGCATTATCTGAAGATAAGAAAAAAATAAAAGACATT
GlyGlyLysIleTyrIleValAspGlyAspIleThrLysHisIleSerLeuGluAlaLeuSerGluAspLysLysLysIleLysAspIle

     910       920       930       940       950       960       970       980       990
TATGGGAAAGATGCTTTATTACATGAACATTATGTATATGCAAAAGAAGGATATGAAGCCGTACTTGTAATCCAATCTTCGGAAGATTAT
TyrGlyLysAspAlaLeuLeuHisGluHisTyrValTyrAlaLysGluGlyTyrGluProValLeuValIleGlnSerSerGluAspTyr

    1000      1010      1020      1030      1040      1050      1060      1070      1080
GTAGAAAATACTGAAAAGGCACTGAACGTTTATTATGAAATAGGTAAGATATTATCAAGGGATATTTTAAGTAAAATTAATCAACCATAT
ValGluAsnThrGluLysAlaLeuAsnValTyrTyrGluIleGlyLysIleLeuSerArgAspIleLeuSerLysIleAsnGlnProTyr

    1090      1100      1110      1120      1130      1140      1150      1160      1170
CAGAAATTTTTAGATGTATTAAATACCATTAAAAATGCATCTGATTCAGATGGACAAGATCTTTTATTTACTAATCAGCTTAAGGAACAT
GlnLysPheLeuAspValLeuAsnThrIleLysAsnAlaSerAspSerAspGlyGlnAspLeuLeuPheThrAsnGlnLeuLysGluHis
```

```
      1180       1190       1200       1210       1220       1230       1240       1250       1260
CCCACAGACTTTTCTGTAGAATTCTTGGAACAAAATAGCAATGAGGTACAAGAAGTATTTGCGAAAGCTTTTGCATATTATATCGAGCCA
ProThrAspPheSerValGluPheLeuGluGlnAsnSerAsnGluValGlnGluValPheAlaLysAlaPheAlaTyrTyrIleGluPro

      1270       1280       1290       1300       1310       1320       1330       1340       1350
CAGCATCGTGATGTTTTACAGCTTTATGCACCGGAAGCTTTTAATTACATGGATAAATTTAACGAACAAGAAATAAATCTATCCTTGGAA
GlnHisArgAspValLeuGlnLeuTyrAlaProGluAlaPheAsnTyrMetAspLysPheAsnGluGlnGluIleAsnLeuSerLeuGlu

      1360       1370       1380       1390       1400       1410       1420       1430       1440
GAACTTAAAGATCAACGGATGCTGTCAAGATATGAAAAATGGGAAAAGATAAAACAGCACTATCAACACTGGAGCGATTCTTTATCTGAA
GluLeuLysAspGlnArgMetLeuSerArgTyrGluLysTrpGluLysIleLysGlnHisTyrGlnHisTrpSerAspSerLeuSerGlu

      1450       1460       1470       1480       1490       1500       1510       1520       1530
GAAGGAAGAGGACTTTTAAAAAAGCTGCAGATTCCTATTGAGCCAAAGAAAGATGACATAATTCATTCTTTATCTGAAGAAGAAAAAGAG
GluGlyArgGlyLeuLeuLysLysLeuGlnIleProIleGluProLysLysAspAspIleIleHisSerLeuSerGlnGluGluLysGlu

      1540       1550       1560       1570       1580       1590       1600       1610       1620
CTTCTAAAAAGAATACAAATTGATAGTAGTGATTTTTTATCTACTGAGGAAAAAGAGTTTTTAAAAAAGCTACAAATTGATATTCGTGAT
LeuLeuLysArgIleGlnIleAspSerSerAspPheLeuSerThrGluGluLysGluPheLeuLysLysLeuGlnIleAspIleArgAsp

      1630       1640       1650       1660       1670       1680       1690       1700       1710
TCTTTATCTGAAGAAGAAAAAGAGCTTTTAAATAGAATACAGGTGGATAGTAGTAATCCTTTATCTGAAAAACAAAAAGAGTTTTTAAAA
SerLeuSerGluGluGluLysGluLeuLeuAsnArgIleGlnValAspSerSerAsnProLeuSerGluLysGluLysGluPheLeuLys

      1720       1730       1740       1750       1760       1770       1780       1790       1800
AAGCTGAAACTTGATATTCAACCATATGATATTAATGAAAGGTTGCAAGATACAGCAGCGTTAATTGATAGTCCGTCAATTAATCTTGAT
LysLeuLysLeuAspIleGlnProTyrAspIleAsnGlnArgLeuGlnAspThrGlyGlyLeuIleAspSerProSerIleAsnLeuAsp

      1810       1820       1830       1840       1850       1860       1870       1880       1890
GTAAGAAAGCAGTATAAAAGGGATATTCAAAATATTGATGCTTTATTACATCAATCCATTGGAAGTACCTTGTACAATAAAATTTATTTTG
ValArgLysGlnTyrLysArgAspIleGlnAsnIleAspAlaLeuLeuHisGlnSerIleGlySerThrLeuTyrAsnLysIleTyrLeu

      1900       1910       1920       1930       1940       1950       1960       1970       1980
TATGAAAATATGAATATCAATAACCTTACAGCAACCCTAGGTGCGGATTTAGTTGATTCCACTGATAATACTAAAATTAATAGAGGTATT
TyrGluAsnMetAsnIleAsnAsnLeuThrAlaThrLeuGlyAlaAspLeuValAspSerThrAspAsnThrLysIleAsnArgGlyIle

      1990       2000       2010       2020       2030       2040       2050       2060       2070
TTCAATGAATTCAAAAAAAATTTCAAATATAGTATTTCTAGTAACTATATGATTGTTGATATAAATGAAAGGCCTGCATTAGATAATGAG
PheAsnGluPheLysLysAsnPheLysTyrSerIleSerSerAsnTyrMetIleValAspIleAsnGluArgProAlaLeuAspAsnGlu

      2080       2090       2100       2110       2120       2130       2140       2150       2160
CGTTTGAAATGGAGAATCCAATTATCACCAGATACTCGAGCAGCATATTTAGAAAATGGAAAGCTTATATTAGAAAGAAACATCCGTCTG
ArgLeuLysTrpArgIleGlnLeuSerProAspThrArgAlaGlyTyrLeuGluAsnGlyLysLeuIleLeuGlnArgAsnIleGlyLeu

      2170       2180       2190       2200       2210       2220       2230       2240       2250
GAAATAAAGGATGTACAAATAATTAAGCAATCCGAAAAAGAATATATAAGGATTGATGCGAAAGTAGTGCCAAAGAGTAAAATAGATACA
GluIleLysAspValGlnIleIleLysGlnSerGluLysGluTyrIleArgIleAspAlaLysValValProLysSerLysIleAspThr

      2260       2270       2280       2290       2300       2310       2320       2330       2340
AAAATTCAAGAAGCACAGTTAAATATAAATCAGGAATGGAATAAAGCATTAGGGTTACCAAAATATACAAAGCTTATTACATTCAACGTG
LysIleGlnGluAlaGlnLeuAsnIleAsnGlnGluTrpAsnLysAlaLeuGlyLeuProLysTyrThrLysLeuIleThrPheAsnVal
```

```
      2350       2360       2370       2380       2390       2400       2410       2420       2430
GATAATAGATATGCATCCAATATTGTAGAAAGTGCTTATTTAATATTGAATGAATGGAAAAATAATATTCAAAGTGATCTTATAAAAAAG
HisAsnArgTyrAlaSerAsnIleValGluSerAlaTyrLeuIleLeuAsnGluTrpLysAsnAsnIleGlnSerAspLeuIleLysLys

      2440       2450       2460       2470       2480       2490       2500       2510       2520
GTAACAAATTACTTAGTTGATGGTAATGGAAGATTTGTTTTTACCGATATTACTCTCCCTAATATAGCTGAACAATATACACATCAAGAT
ValThrAsnTyrLeuValAspGlyAsnGlyArgPheValPheThrAspIleThrLeuProAsnIleAlaGluGlnTyrThrHisGlnAsp

      2530       2540       2550       2560       2570       2580       2590       2600       2610
GAGATATATGAGCAAGTTCATTCAAAACGGTTATATGTTCCAGAATCCCGTTCTATATTACTCCATGGACCTTCAAAAGGTGTAGAATTA
GluIleTyrGluGlnValHisSerLysGlyLeuTyrValProGluSerArgSerIleLeuLeuHisGlyProSerLysGlyValGluLeu

      2620       2630       2640       2650       2660       2670       2680       2690       2700
AGGAATGATAGTGAGGGTTTTATACACGAATTTGGACATGCTGTGGATGATTATGCTGGATATCTATTAGATAAGAACCAATCTGATTTA
ArgAsnAspSerGluGlyPheIleHisGluPheGlyHisAlaValAspAspTyrAlaGlyTyrLeuLeuAspLysAsnGlnSerAspLeu

      2710       2720       2730       2740       2750       2760       2770       2780       2790
GTTACAAATTCTAAAAAATTCATTGATTTTTTAAGGAAGAAGGGAGTAATTTAACTTCGTATGGCAGAACAAATGAAGCGGAATTTTTT
ValThrAsnSerLysLysPheIleAspIlePheLysGluGluGlySerAsnLeuThrSerTyrGlyArgThrAsnGluAlaGluPhePhe

      2800       2810       2820       2830       2840       2850       2860       2870       2880
GCAGAAGCCTTTAGGTTAATGCATTCTACGGACCATGCTGAAGGTTTAAAAGTTCAAAAAAATGCTCCGAAAACTTTCCAATTTATTAAC
AlaGluAlaPheArgLeuMetHisSerThrAspHisAlaGluArgLeuLysValGlnLysAsnAlaProLysThrPheGlnPheIleAsn

      2890       2900       2910       2920       2930       2940       2950       2960       2970
GATCAGATTAAGTTCATTATTAACTCATAAGTAATGTATTAAAAATTTTCAAATGGATTTAATAATAATAATAATAATAATAACGGG
AspGlnIleLysPheIleIleAsnSer

      2980       2990       3000       3010       3020       3030       3040       3050       3060
ACCAGCCATTATGAAGCAACTAATTCTAGACTTGATAGTAATTCTTGGGAAGCACCAGATAGTGTAAAAGGTGGCATTGCCAGAATGATA


      3070       3080       3090       3100       3110       3120       3130       3140       3150
TTTTATGTGTTCGTTAGATATGAAGGCAAAAACAATGATCCTGACCTAGAAGTTAATGATAATGTTATTAATAATTTAATGCCTTTTATA

      3160       3170       3180       3190       3200       3210       3220       3230       3240
GGAATATTAGTAAAAGTGCCGAAAAGATCCTGTTGCAAAGCTTTTAAAGAACATATTATTCTATCAAGTGGCTGTATATTTTGTGTAATT

      3250       3260       3270       3230       3290
TTCAATAAATTTTGTAATTAAGCATACGTCAAAAAACCGAAATCTGAGCTC
                                                 SstI
```

APPENDIX III. LF amino acid sequence

(29 aa signal peptide) ↓-Start of mature LF (780 aa)

```
  1 MNIKKEFIKVISMSCLVTAITLSGPVFIPLVQGAGGHGDVGMHVKEKEKNKDENKRKDEERNKTQEEHLK

 71 EIMKHIVKIEVKGEEAVKKEAAEKLLEKVPSDVLEMYKAIGGKIYIVDGDITKHISLEALSEDKKKIKDI

141 YGKDALLHEHYVYAKEGYEPVLVIQSSEDYVENTEKALNVYYEIGKILSRDILSKINQPYQKFLDVLNTI
                                                  #1

211 KNASDSDGQDLLFTNQLKEHPTDFSVEFLEQNSNEVQEVFAKAFAYYIEPQHRDVLQLYAPEAFNYMDKF
      #2                                                                  #3

281 NEQEINLSLEELKDQRMLSRYEKWEKIKQHYQHWSDSLSEEGRGLLKKLQIPIEPKKDDIIHSLSQEEKE

351 LLKRIQIDSSDFLSTEEKEFLKKLQIDIRDSLSEEEKELLNRIQVDSSNPLSEKEKEFLKKLKLDIQPYD

421 INQRLQDIGGLIDSPSINLDVRKQYKRDIQNIDALLHQSIGSTLYNKIYLYENMNINNLTATLGADLVDS

491 TDNTKINRGIFNEFKKNFKYSISSNYMIVDINERPALDNERLKWRIQLSPDTRAGYLENGKLILQRNIGL

561 EIKDVQIIKQSEKEYIRIDAKVVPKSKIDTKIQEAQLNINQEWNKALGLPKYTKLITFNVHNRYASNIVE

631 SAYLILNEWKNNIQSDLIKKVTNYLVDGNGRFVFTDITLPNIAEQYTHQDEIYEQVHSKGLYVPESRSIL

701 LHGPSKGVELRNDSEGFIHEFGHAVDDYAGYLLDKNQSDLVTNSKKFIDIFKEEGSNLTSYGRTNEAEFF

771 AEAFRLMHSTDHAERLKVQKNAPKTFQFINDQIKFIINS
```

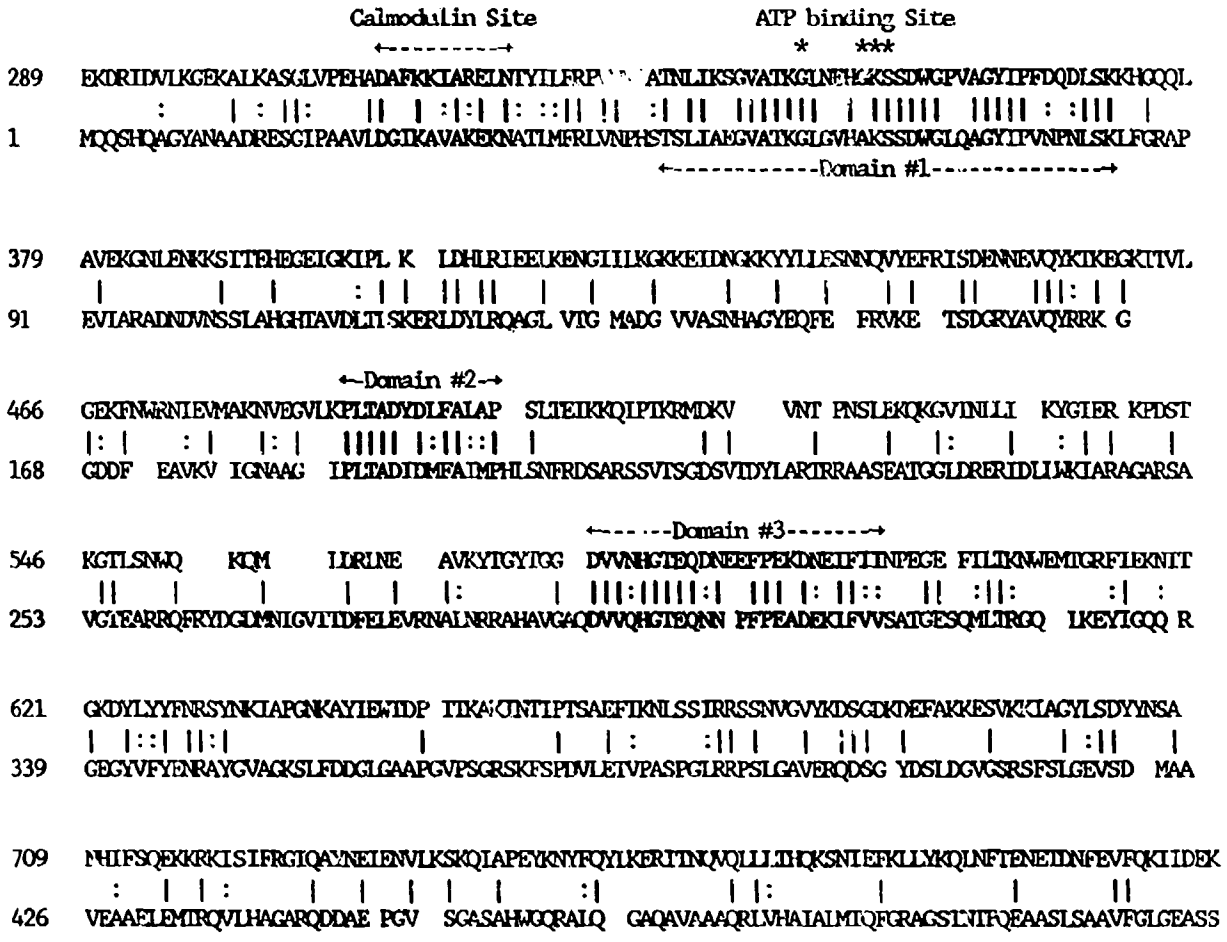
****

The sequence contains 809 amino acids (M_r 93,798):

Ala (A)	34	Leu (L)	80
Arg (R)	27	Lys (K)	86
Asn (N)	54	Met (M)	10
Asp (D)	55	Phe (F)	29
Cys (C)	1	Pro (P)	21
Gln (Q)	41	Ser (S)	54
Glu (E)	79	Thr (T)	28
Gly (G)	35	Trp (W)	5
His (H)	21	Tyr (Y)	35
Ile (I)	74	Val (V)	40

Acidic	(Asp + Glu)	134
Basic	(Arg + Lys)	113
Aromatic	(Phe + Trp + Tyr)	69
Hydrophobic	(Aromatic + Ile + Leu + Met + Val)	273

APPENDIX IV. Homology Comparison between EF and pertussis cyclase.

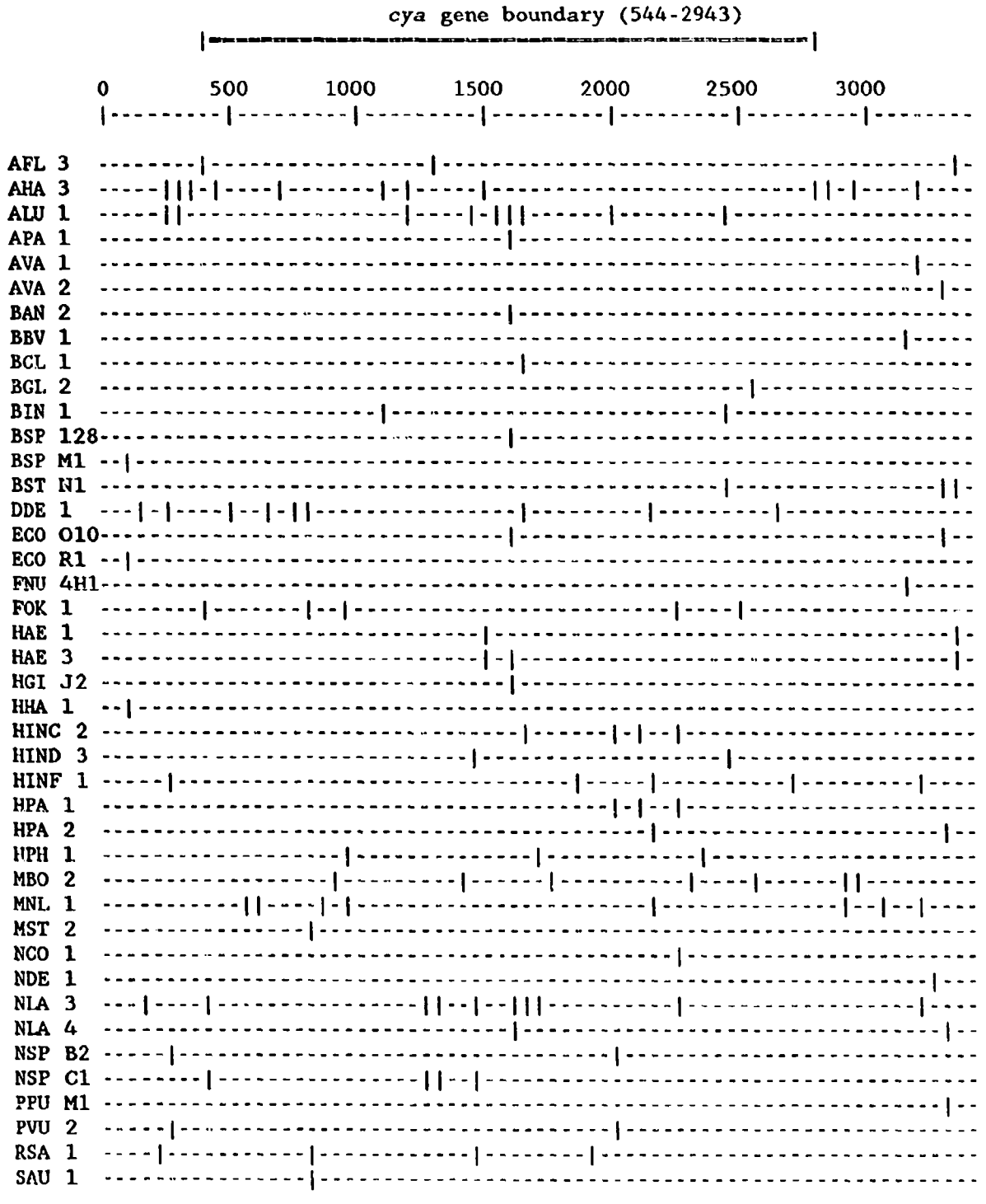


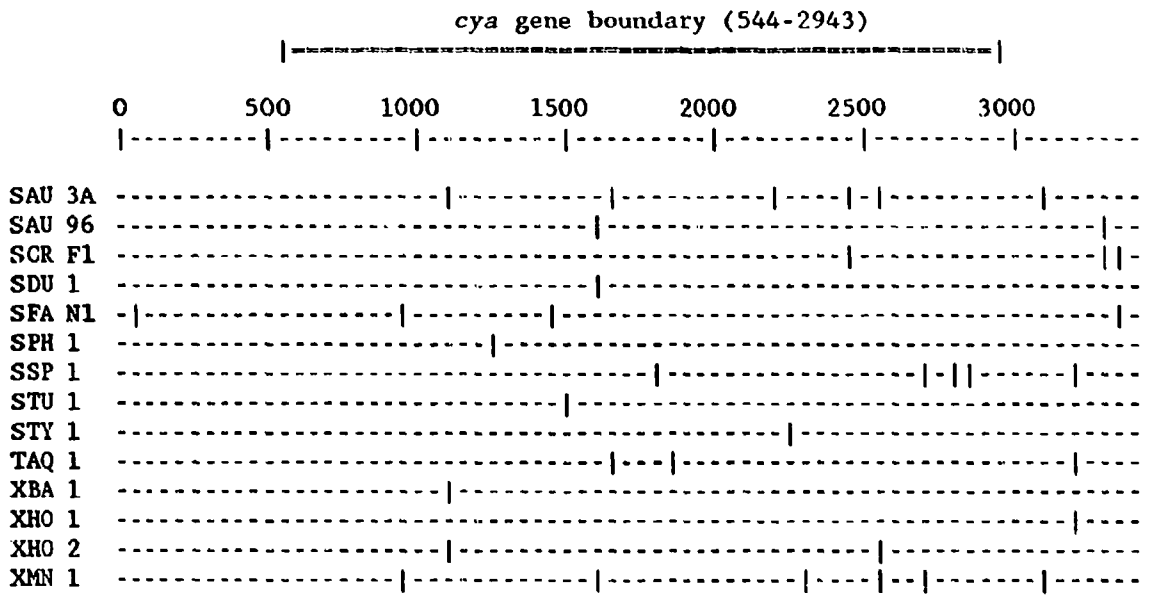
1. Domains #1, #2 and #3 represent three highly conserved amino acid domains in EF (top line of each pair) and the pertussis cyclase (bottom line in each pair).

2. The numbers to the left of each line indicates the amino acid position for EF-precursor or the pertussis cyclase.

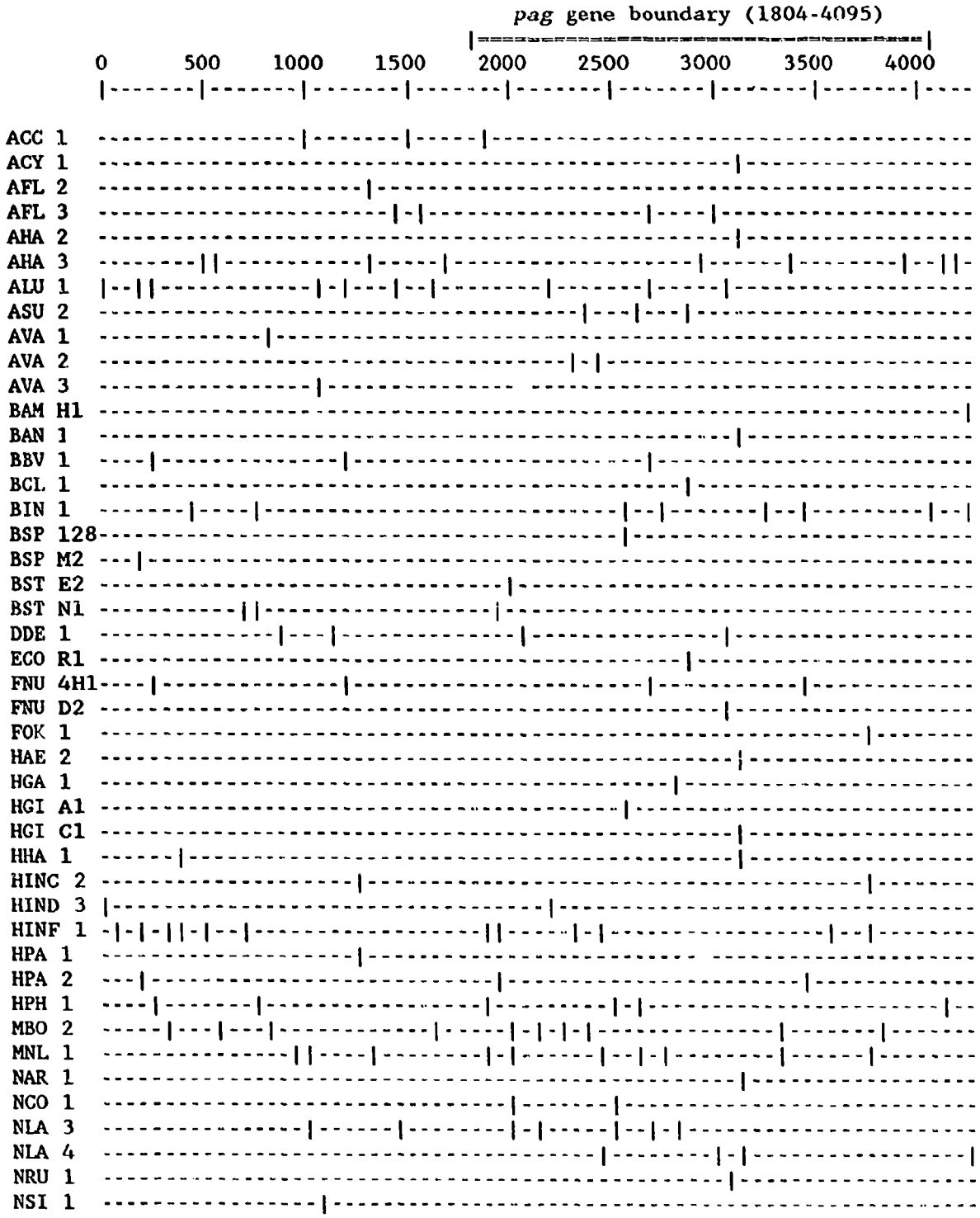
3. The asterisks (*) indicate the consensus sequences for the ATP binding site for EF and the pertussis cyclase.

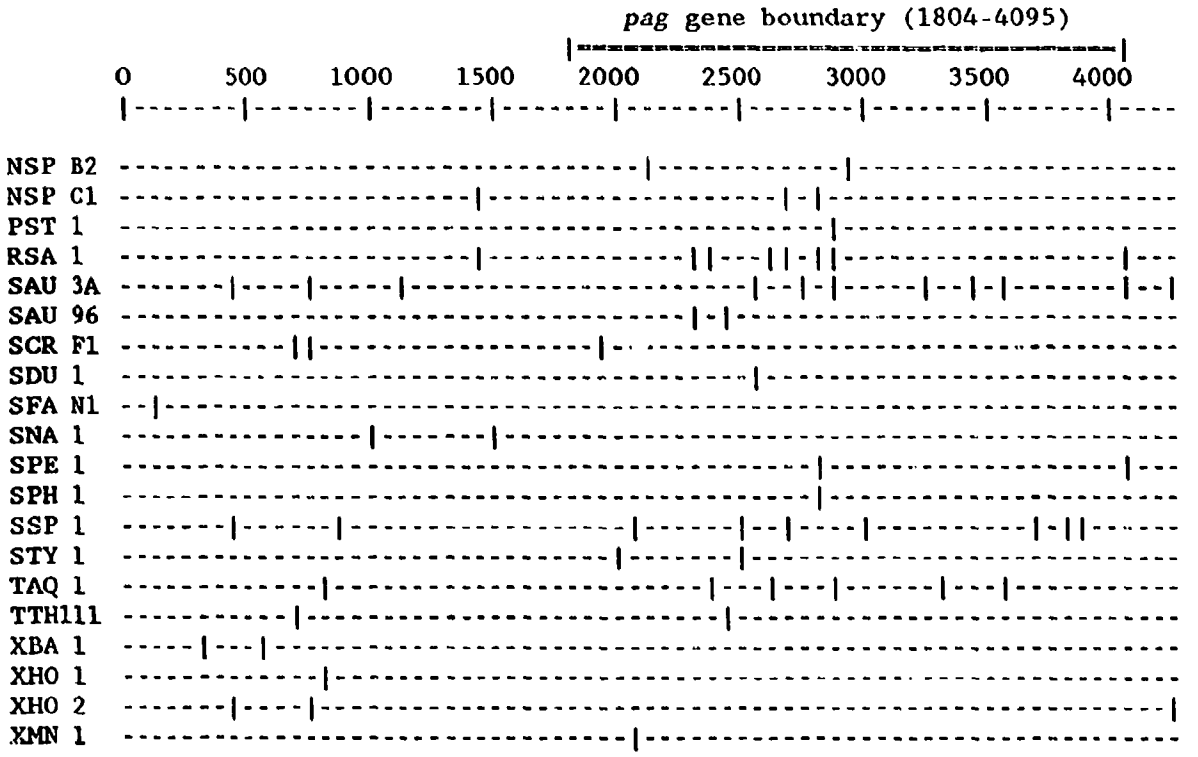
APPENDIX V: Restriction enzyme cleavage sites for EF gene.



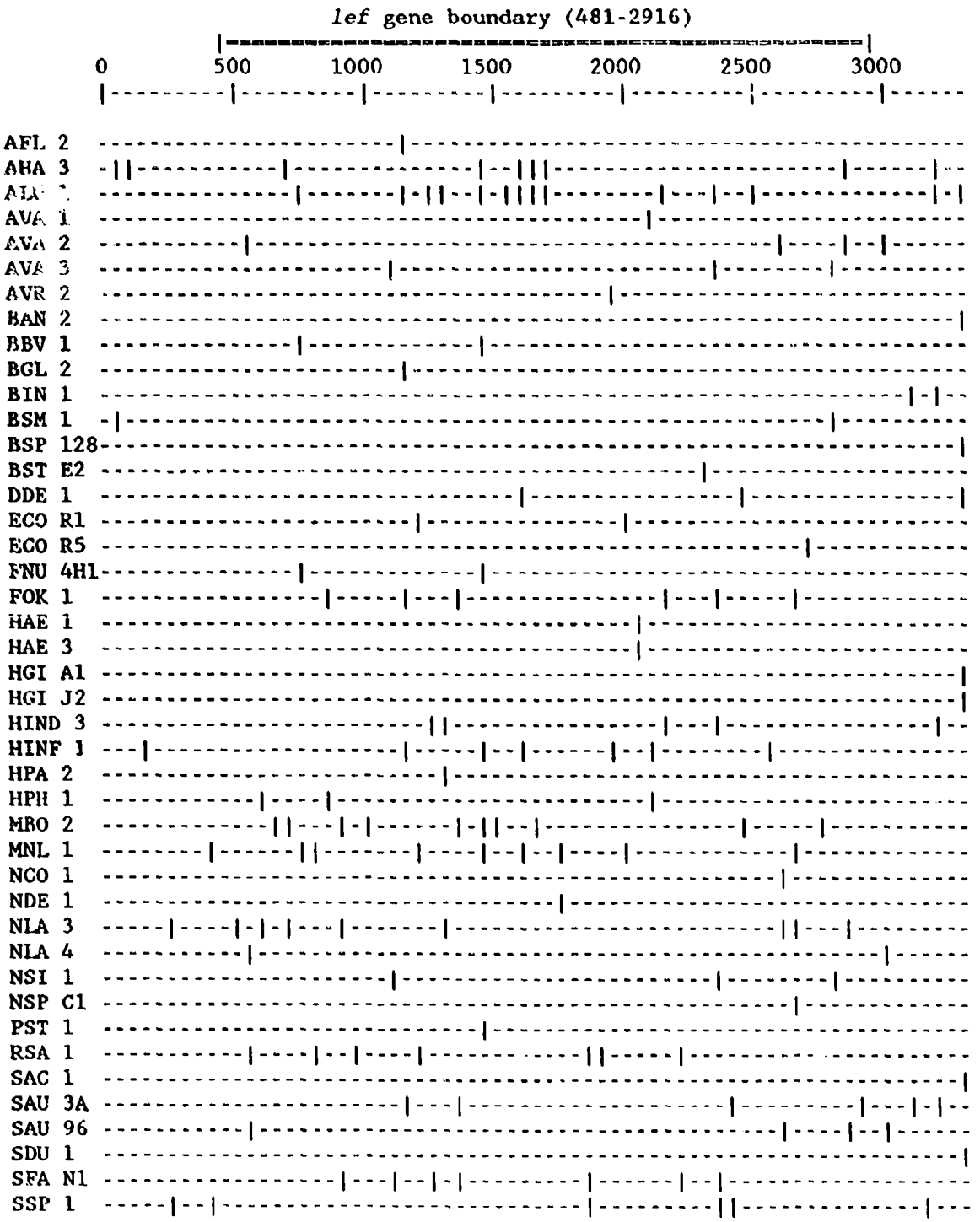

cya gene boundary (544-2943)
0
500
1000
1500
2000
2500
3000
SAU 3A
SAU 96
SCR F1
SDU 1
SFA N1
SPH 1
SSP 1
STU 1
STY 1
TAQ 1
XBA 1
XHO 1
XHO 2
XMN 1

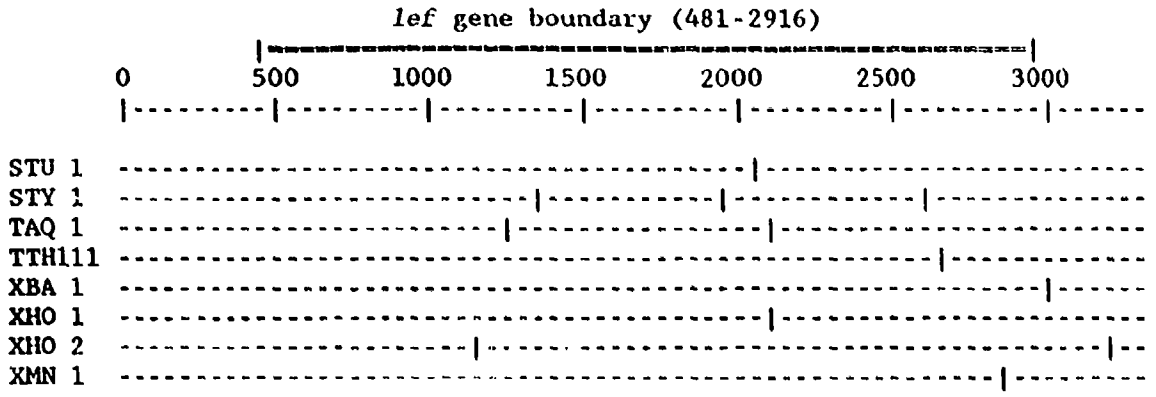
APPENDIX VI: Restriction enzyme cleavage site for PA gene




pag gene boundary (1804-4095)
0
500
1000
1500
2000
2500
3000
3500
4000
NSP B2
NSP C1
PST 1
RSA 1
SAU 3A
SAU 96
SCR F1
SDU 1
SFA N1
SNA 1
SPE 1
SPH 1
SSP 1
STY 1
TAQ 1
TTH111
XBA 1
XHO 1
XHO 2
XMN 1

APPENDIX VII: Restriction enzyme cleavage site for LF gene




lef gene boundary (481-2916)
0
500
1000
1500
2000
2500
3000
STU 1
STY 1
TAQ 1
TTH111
XBA 1
XHO 1
XHO 2
XMN 1

DISTRIBUTION LIST

5 copies — Commander
US Army Medical Research Institute of
Infectious Diseases
ATTN: SGRD-UIZ-M
Fort Detrick, Frederick, MD 21701-5011

1 copy — Commander
US Army Medical Research and Development Command
ATTN: SGRD-RMI-S
Fort Detrick, Frederick, MD 21701-5012

2 copies — Defense Technical Information Center (DTIC)
ATTN: DTIC-DDAC
Cameron Station
Alexandria, VA 22304-6145

1 copy — Dean
School of Medicine
Uniformed Services University of the
Health Sciences
4301 Jones Bridge Road
Bethesda, MD 20814-4799

1 copy — Commandant
Academy of Health Sciences, US Army
ATTN: AHS-CDM
Fort Sam Houston, TX 78234-6100